Sunspot positions and sizes for 1825-1867 from the observations by Samuel Heinrich Schwabe

Arlt, R.; Leussu, R.; Giese, N.; Mursula, K.; Usoskin, I. G.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stt961

Publication date:
2013

Citation for published version (APA):
Arlt, R., Leussu, R., Giese, N., Mursula, K., & Usoskin, I. G. (2013). Sunspot positions and sizes for 1825-1867 from the observations by Samuel Heinrich Schwabe. Monthly Notices of the Royal Astronomical Society, 433(4), 3165-3172. https://doi.org/10.1093/mnras/stt961

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **433,** 3165–3172 (2013) doi:10.1093/mnras/stt961
Advance Access publication 2013 June 21

Sunspot positions and sizes for 1825–1867 from the observations by Samuel Heinrich Schwabe

R. Arlt,[1]★ R. Leussu,[2] N. Giese,[1,3] K. Mursula[2] and I. G. Usoskin[2,4]

[1] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[2] *Department of Physics, Centre of Excellence in Research, University of Oulu, PO Box 3000, 90014 Oulu, Finland*
[3] *Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, the Netherlands*
[4] *Sodankylä Observatory (Oulu unit), University of Oulu, PO Box 3000, 90014 Oulu, Finland*

Accepted 2013 May 30. Received 2013 May 29; in original form 2013 April 19

ABSTRACT

Samuel Heinrich Schwabe made 8486 drawings of the solar disc with sunspots in the period from 1825 November 5 to 1867 December 29. We have measured sunspot sizes and heliographic positions on digitized images of these drawings. A total of about 135 000 measurements of individual sunspots are available in a data base. Positions are accurate to about 5 per cent of the solar radius or to about 3° in heliographic coordinates in the solar-disc centre. Sizes were given in 12 classes as estimated visually with circular cursor shapes on the screen. Most of the drawings show a coordinate grid aligned with the celestial coordinate system. A subset of 1168 drawings have no indication of their orientation. We have used a Bayesian estimator to infer the orientations of the drawings as well as the average heliographic spot positions from a chain of drawings of several days, using the rotation profile of the present Sun. The data base also includes all information available from Schwabe on spotless days.

Key words: Sun: activity – sunspots.

1 INTRODUCTION

It is desirable to compile a time series of individual sunspot positions going back to the time when telescopes were first used to observe them. Such a time series will contain an enormous amount of features of great importance for the solar dynamo and the theory of magnetic flux emergence at the solar surface. A list of existing time series was compiled by Lefevre & Clette (2012). Data of individual spots were not available for the period before the Kodaikanal data starting in 1906 until the analyses of the Staudacher drawings by Arlt (2009) covering 1749–1799 and the Zucconi drawings by Cristo, Vaquero & Sánchez-Bajo (2011) covering 1754–1760.

The first paper (Arlt 2011, hereafter Paper I) focused on the inventory and description of the digitization of the historical sunspot drawings by Samuel Heinrich Schwabe made in the period of 1825–1867. The majority of drawings were made with a high-quality Fraunhofer refractor of 3.5 feet focal length.

The full set of 8486 full-disc drawings has now been fully measured. The method of measurements will be described in Section 2 while the resulting spot distribution and the possible sources of errors will be discussed in Section 3. The analysis aims at the full exploitation of the drawings by providing positional information of each individual sunspot together with its size. Unfortunately, the data set by the Royal Greenwich Observatory (RGO) and its continuation by the US Air Force and the National Oceanic and Atmospheric Administration (USAF/NOAA) only provides the average group positions and the total areas of the groups.[1] Information like the size distribution of sunspots and the tilt angles and polarity separations of bipolar regions is only preserved if the individual spots are stored in the data set, however.

The Schwabe data are also superior to the ones by Carrington (1853–1861; cf. Lepshokov, Tlatov & Vasil'eva 2012 for a recent analysis) and Spörer (1861–1894; recent analysis by Diercke, Arlt & Denker 2012), which only report about sunspot groups at a certain instance when they were near the central meridian. The Schwabe data contain the full evolution of sunspot groups crossing the visible solar disc.

2 METHODS OF MEASUREMENTS

2.1 Heliographic coordinate system

For all images possessing a horizontal reference line, we assumed that the line is parallel to the celestial equator (cf. Paper I). The position angle and tip angle of the heliographic coordinate system are obtained from the JPL HORIZONS ephemeris generator.[2] We used the geographical coordinates of the observing location in the town of Dessau, Germany, and generated a list of these quantities in 6 h intervals for the entire period of 1825–1867. The quantities for

---

★ E-mail: rarlt@aip.de

[1] Hathaway, http://solarscience.msfc.nasa.gov/greenwch.shtml

[2] http://ssd.jpl.nasa.gov/horizons.cgi

times in between two output lines were interpolated linearly. The documentation of the HORIZONS ephemeris service states that the position angle is the 'target's North Pole position angle (counter-clockwise with respect to direction of true-of-date celestial north)'. It is reasonable to assume that Schwabe used the local sky rotation to adjust his telescope to the north. He must thus have arrived nearly at a 'true-of-date' celestial north. The actual orientation of the solar-disc drawing comes from the cross-hairs used in the eyepiece. Schwabe did not report on how he adjusted the eyepiece (rotation may have easily been possible). Throughout the vast majority of observations the alignment is amazingly consistent but, as we will see later, there are a few short periods when the eyepiece was apparently misaligned.

For all observing days with drawings, the actual solar disc was extracted from the digitized image by four mouse clicks on the left, right, lower and upper limbs of the circle, where the middle of the pencil stroke width was chosen. This way, slight ellipticities are also allowed thereby, although only in the vertical or horizontal directions and not at an arbitrary angle. This turned out to be a reasonable choice, since ellipticities mainly come from the fact that the paper may not have been entirely flat when photographed, producing a prolateness of the circles.

If a horizontal line is available in the image, two clicks near the left end and near the right end of the line define the position angle of the celestial equator in the image. Again, the middle of the pencil stroke width was chosen visually. The position angle of the solar equator is then added to this orientation, and the actual heliographic grid is superimposed to the image.

In some cases, the main vertical line is not perpendicular to the horizontal one. We are applying a special transformation to the Cartesian coordinates of the measurements, as described in Section 2.4. The various tools for the measurements were written in the Interactive Data Language (IDL).

2.2 Method for unoriented drawings

There is a set of 1168 drawings which do not show a coordinate system, mostly in the period of mid-1826 to 1830. Since there are often sequences of days for which the drawings have a number of sunspots in common, we can use the rotation of the Sun to find the probable position angles of the heliographic coordinate systems. We assume the sidereal rotation profile obtained from average sunspot group positions by Balthasar, Vázquez & Wöhl (1986) and use the numerical values

$$\Omega(b) = 14\overset{\circ}{.}551\,\mathrm{d}^{-1} - 2\overset{\circ}{.}87\,\mathrm{d}^{-1}\sin^2 b \qquad (1)$$

for the angular velocity Ω, where b is the heliographic latitude. We actually need the synodic rotation rate for our purposes which is obtained from solving Kepler's equation for the eccentric anomaly of the Earth at each instance it is needed, using an eccentricity of $e = 0.016\,87$ and a rotation period of $P_{\mathrm{rot}} = 365.242\,198\,79\,\mathrm{d} - 6.14 \times 10^{-1}(\mathrm{JD} - 241\,5019)/36\,525$ (Newcomb 1898). Note that the use of the solar rotation profile implies that we cannot use the resulting sunspot positions directly for the determination of the differential rotation of the Sun later on, since they are not independent of the rotation profile.

A Bayesian parameter estimation is employed to obtain the position angles and average sunspot positions. We start with looking at n_d drawings and associating n_s sunspots with each other, which are visible in all these drawings. Given the two coordinates of each spot, these combinations deliver $N = 2n_\mathrm{s}n_\mathrm{d}$ measurements. The unknowns are the heliographic coordinates of the spots, l_i and b_i, where $i = 1, \ldots, n_\mathrm{s}$ counts the spots and the position angles p_j of the drawings where $j = 1, \ldots, n_\mathrm{d}$ counts the days. We are thus faced with $M = 2n_\mathrm{s} + n_\mathrm{d}$ free parameters. For three days with three common spots, we have $N = 18$ measurements and $M = 9$ unknowns, for example, while two days with two spots deliver only $N = 8$ and $M = 6$. Note that there may be two or three days between two adjacent drawings in a sequence.

Formally, there is another parameter which we either have to determine beforehand or keep as a free parameter. It is the measurement error of Schwabe's plots. It is reasonable to assume that these errors roughly form a Gaussian distribution. Deviations from Gaussian distributions may only be expected for spots very near the solar limb, but for the majority of spots, Gaussian will be a good approximation, and we assume that there is a single standard deviation σ describing the distribution. Allowing σ to be a free parameter was considered, but turned out to be impractical since the model then obtains excessive freedom to assume that the spots are in the wrong place and yield very odd combinations of latitudes and position angles at high likelihood. The value of σ was thus estimated from a number of chains with high n_d and high n_s using the residuals. These should be identical to the plotting errors only for infinitely large n_d and n_s, an exactly known rotation profile and the assumption of zero proper motion of the spots. As a compromise, we chose chains of five drawings having 2–4 spots in common and kept σ as a free parameter. From this set of 15 sample chains (i.e. 75 drawings) in 1827 and 1828, we obtained an average $\sigma = 0.05$ of the solar-disc radius. We used this value of σ in all actual determinations of position angles of drawings where no coordinate system was given by Schwabe. Note also the additional remarks about the accuracy in Section 4.

Bayesian inference is based on the distribution of probability density over the entire parameter space. (We will often use the term 'probability distribution', but actually the probability density is meant.) Every combination of parameters, given the model differential rotation, is tested on its likelihood to have created the data. Since this is too expensive computationally, we are employing Monte Carlo Markov chains which explore the parameter space very efficiently, without wasting the computing time in regions of very low probability density but without being limited to local maxima either. The parameter space for the determination of orientations from one chain is binned into 2048^M bins for which the number of passages of the Markov chains is counted. After normalization, these counts give the probability density distribution. The posterior distribution for a given individual parameter is obtained by marginalization over all the other parameters.

One often has several options of combining consecutive drawings into a chain that is analysed by the Bayesian estimator. It is of course not a matter of the residuals to tell which combination is best, since the residuals always improve when the number of free parameters approaches the number of measurements. We will denote the combinations by $n_\mathrm{d}/n_\mathrm{s}$ in the following.

The suitability of combinations of drawings for the determination of the orientations is not easily quantified. The Bayesian information criterion (BIC, also called Schwartz criterion) is one guess for the trade-off between keeping the residuals as well as the number of free parameters low. It does not, however, take into account the distribution of the spots over the solar disc which may vary from very suitable to almost degenerate. We computed about 20 test cases to obtain an idea of good and bad distributions. Based on the BIC and this experience, we start from the combination 3/3 as the desired one and used a ranking for other combinations to be chosen when 3/3 is not possible. The ranking with descending 'priority' is the following: 3/3, 2/6, 2/5, 3/4, 4/2, 2/4, 2/3 and 3/2 (turned out to be equal in suitability), 2/2, 6/1. Rare occasion with many common spots are not listed here, but the first three combinations indicate already, that it is better to use three drawings with few spots than a pair of drawings with many spots. Since there is considerably more rotational displacement with three drawings, the third drawing fixes the spot positions very well and always produces highly plausible results.

A subjective quality flag is given to the spots of a given day. All drawings with a pencil grid obtain a quality flag of $Q = 1$. Positions derived from the rotational matching of two or more consecutive drawings may obtain $Q = 1$, 2 or 3. Our rules to assign the different quality flags are as follows: if the probability distribution of one of the free parameters has a full confidence interval of more than 40°, but the distributions are not skewed, we assign a quality of $Q = 2$. If the distributions are additionally slightly skewed so that the average parameter is different from the mode value by up to 20°, we use $Q = 3$. A subjective estimate of the quality is given with values from 1 (highest quality) to 3 (lowest quality). Drawings delivering very skewed or double-peaked distributions are discarded and the quality estimate is set to 4 (see Section 2.7). We did not derive any sunspot positions for those days. We store the sunspots of discarded drawings and fill their positions with NaN. Yet their sizes are available and useful and are stored in the spot file along with group designations.

It is of great advantage to know the full distribution of the probability density as compared to minimization procedures for, e.g., χ^2. Such searches are not aware of additional minima and may even miss the global minimum entirely.

2.3 On-screen measurements

We used a set of circular mouse cursor shapes with different diameters to estimate the sizes and positions of the sunspots. For all spots showing a penumbra, only the umbral size was measured. This is because the open circles were often drawn by Schwabe to indicate the presence of penumbrae. While the umbrae pencil dates are clearly drawn with the intention to distinguish different sizes, the penumbrae show less carefulness since they are all of very similar size. Additionally, Schwabe's penumbrae show little foreshortening near the solar limb (see group 68 in Fig. 1). We leave it to future scrutinization which may or may not show the scientific usefulness of the penumbral sizes drawn by Schwabe.

A total of 12 size steps of a circular cursor were used running from an area of 5 square pixels to 364 square pixels (Table 1) including the borders. We always used the largest possible circular cursor for which the boundary of the circle was contained *within* the umbral area, if the umbra was circular. Non-circular spots can only be approximately matched with these cursor masks, of course. Note that the pencil dots have a certain minimum size which did not require the use of 1 square-pixel areas. The total area of the solar disc is 708 822 pixels. A single square pixel corresponds to 1.4 millionths of the disc. The smallest areas measured here are 7 millionths of the solar disc. An alternative way of estimating the areas was given by Cristo et al. (2011). In their work, the umbral areas were derived for Zucconi's observations in 1754–1760 in a semi-automatic black-pixel-finding algorithm which can deal with almost arbitrary sunspot shapes. Because of the lower and varying contrast in Schwabe's pencil drawings, this algorithm would be more difficult to apply in our case, and was not employed.

Figure 1. Example drawing of 1836 April 11 with penumbrae. Most of Schwabe's drawings are made in this style.

Table 1. Cursor sizes and corresponding areas in square pixels.

Size	Area
1	5
2	9
3	21
4	37
5	69
6	97
7	145
8	185
9	206
10	270
11	308
12	364

Before 1831, Schwabe did not distinguish umbra and penumbra in his drawings. The first full-disc drawing with distinguished penumbrae is from 1831 January 06. At the same time, Schwabe stopped drawing magnifications of sunspot groups besides the full-disc drawings on a regular basis and did so only for spectacular groups or interesting observational facts he wanted to emphasize. We will have to choose an appropriate calibration for the sunspot areas in order to obtain a consistent data set.

We did not contemplate using elliptical cursor shapes for foreshortened sunspots near the solar limb. The cursor size was chosen visually as to approximate the roughly elliptical shape of the sunspot by a circle of equal area instead, but still referring to the projected sunspot area. The introduction of different ellipticities for different limb distances would have made the measurements considerably more time consuming.

For the sunspot position, the appropriate cursor shape was centred on the pencil dot in the image visually and the position was fixed by a mouse click. We decided to use only the spots visible in the full-disc drawings, delivering a consistent set of spots always drawn at the same scale. Detailed drawings of sunspot groups next to the full-disc drawings were not used despite containing additional fine pores.

All positions were first stored in a momentary reference frame with the 0° meridian running through the centre of the disc (central meridian distance, CMD). If the interpretation of the times given by Schwabe should change, new Carrington longitudes could always be generated from the momentary reference frames.

2.4 Skewed coordinate systems

The main vertical and horizontal lines are not always perfectly perpendicular. In cases where the difference from 90° is more than $1\overset{\circ}{.}66$ (corresponding to roughly half the plotting accuracy – see Section 4), we applied a transformation to the normalized, Cartesian coordinates before we converted them into heliographic ones.

Since it is the lines on paper that have to be drawn anew every day, while the actual cross-hairs in the eyepiece need no re-alignment, we assume that the eyepiece was correct, whereas the drawing was imperfectly made.

When copying the visual information on the spot positions from the eyepiece, the lines were used as references. If the lines on paper differ from the view in the eyepiece, the (additional) plotting error is larger the closer the spot is to one of the reference lines. The spots near any of the lines will be offset by the same amount as the reference line is offset against the real view in the telescope.

Let us consider a polar coordinate system with the intersection between the 'horizontal' and 'vertical' reference lines being the origin. Any spot will appear in a sector between such 'horizontal' and 'vertical' lines. Let ϕ_1 and ϕ_2 be the two angles at which these 'horizontal' and 'vertical' lines are drawn. We also convert the measured Cartesian (x, y) into polar coordinates (r, ϕ) on the solar disc. The angles ϕ_1 and ϕ_2 as well as the Cartesian and polar coordinates are defined in the usual rectangular coordinate system aligned with the image coordinates which is only of auxiliary nature. The situation is depicted in Fig. 2, where the deviation from perpendicularity is exaggerated for clarity. The correct Schwabe system is now positioned in such a way that the new lines have equal angular distances from the plotted 'horizontal' and 'vertical' lines, respectively, and are perpendicular (not plotted in Fig. 2). This angular distance is denoted by α. We correct the spot position by

$$\phi' = \phi + \alpha \left(\frac{2\phi}{\phi_2 - \phi_1} - 1 \right)^q , \qquad (2)$$

where the new location is (r, ϕ'). α is the deviation of the vertical and the horizontal lines from being rectangular, $\alpha = (\phi_2 - \phi_1 - 90^\circ)/2$. The term in parentheses in equation (2) gives numbers between -1 and 1 which are multiplied by the maximum shift which would be necessary if the spot is exactly on one of the wrong axes at ϕ_1 or ϕ_2. The exponent q controls the strength of the re-mapping. A small q causes the re-mapping to be effective over most of the sector between a horizontal and a vertical line. A large q causes the re-mapping to be confined close to the lines while being practically zero in the 'field' between the lines. We used $q = 1$ throughout the analysis.

2.5 Typical problems occurring

All measurements are made manually. This allowed us to interpret what is meant in the drawing at every instance of the process. Some features in the images can mimic sunspots and need to be distinguished.

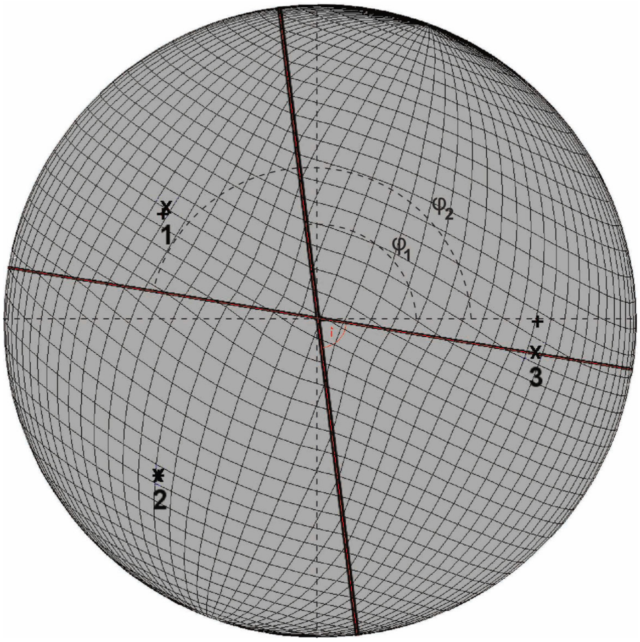

Figure 2. Highly exaggerated test case for the correction of skewed coordinate systems. The *x*-shape symbols represent spots in the drawing, the +− symbols are the corrected positions. The angles ϕ_1 and ϕ_2 are used in equation (2).

(i) Paper defects. They usually have a slightly brownish colour and can be distinguished from pencil-drawn sunspots quite easily.

(ii) Faculae were often marked in the drawings, but of course not as bright features but with weak, often curved pencil strokes. Visual inspection often tells what are faculae in a given group and what are small spots. Faculae without spots (especially near the solar limb) do not have group numbers and can thus be omitted from the measurements. In doubtful cases, the verbal descriptions can be used as they regularly report on the presence of faculae ('Lichtgewölk').

(iii) Dots associated with group numbers. Schwabe often added a dot after the group number to mark it as an ordinal number. Also the number 1 gets a top dot to distinguish it from a simple vertical line. Hence, group number 11 comes along with two additional dots in the drawing. They are drawn in ink and appear darker than the pencil-drawn sunspots.

(iv) Pinholes from the pair of compasses. While the pinhole of the actual drawing is obviously not disturbing as it is passed by the vertical and horizontal reference lines, pinholes from drawings on the back of the paper can appear anywhere in the solar disc. They can be distinguished from spots since they exhibit a raised appearance in contrast to the engraved pencil dots of the real spots.

2.6 Group numbers

Schwabe numbered the groups starting with number 1 each year. A few groups visible already in the previous year carried their numbers into the new year. Schwabe tried to identify groups from previous solar rotations when they became visible again. He mentioned

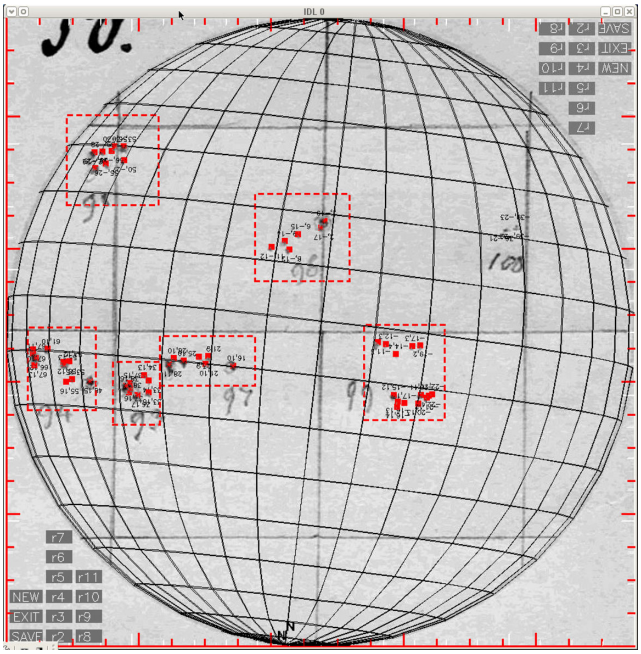

Figure 3. Screen-shot of the group numbering tool for the drawing of 1861 June 22. The actual measurement was made with an image rotated by 180°, since Schwabe's drawings are all upside-down. This is why the coordinates are upside-down, while it is more convenient to use Schwabe's original orientation for reading the group numbers. The picture is a screen-shot from the process of numbering, whence the yet unnumbered group 100.

possible re-apparitions but always assigned new numbers to any group appearing on the eastern limb of the Sun.

We store the group designations for each spot measured. They are not always numbers. In the very beginning, Schwabe used letters. Faculae – most prominently visible near the solar limb – were often referred to by Greek letters. When Schwabe referred to parts of a group in his verbal descriptions, he also used Greek letters very often.

Note that the definition of a group is not necessarily identical to a group definition we would use today. Two bipolar groups at the same heliographic longitude but at slightly different latitudes were most likely classified as a common group although we would separate them as two groups with today's knowledge. Another difficulty arises from the foreshortening when new groups appear near the limb. Schwabe assigned a single group number to some sunspots appearing at the limb, although they turn out to be two or more groups when the full longitudinal extent becomes evident in the middle of the solar disc. An example of the numbering is given in Fig. 3. While the numbering is typically fine, we also see an example (group no. 99) where two groups were combined into one group. Nevertheless, we kept the original group numbers to preserve as much of the historical information of the drawings as possible.

3 DESCRIPTION OF THE DATA FILE

The data are arranged in a format described in Table 2. There is a single blank space between each of the data fields. The first five fields contain the time to which the positions refer. It is fairly certain that the times of observations are mean local times, since Schwabe made efforts to determine culmination times and keep track of deviations of his clocks of the order of seconds. In some cases, the time to which the full-disc drawing refers is ambiguous or missing. When missing, we assumed 12 h local time and set the Timeflag = 0 for these cases. When several times were given ambiguously, we used the most probable time given – in most cases also 12 h as the days before and after are typically stating 12 h clearly for the times of the drawings.

The fields L0 and B0 are the heliographic coordinates of the centre of the Sun as given by the JPL HORIZONS ephemeris service ('Observer sublong and sublat'). The coordinates are for the apparent disc centre as seen from the observing location in Dessau, but the differences to the topocentric coordinates are far below the plotting accuracy (parallax of the order of $0\overset{\circ}{.}002$). L0 and B0 are equal for all spots of a given day, of course. But we give them for every spot to ensure that the conversion to the Carrington frame of heliographic coordinates will be replicable. Note that HORIZONS obtains a zero longitude for the disc centre on 1853 November 9, at $21^{\mathrm{h}}36^{\mathrm{m}}$ UT. Carrington defined the zero-point of his longitude counting on 1853 November 9.

The CMD is the central meridian distance and is a heliographic longitude measured from the central meridian where values west of it (seen on the observer's sky) are positive and values east of it are negative. The direction of measuring longitudes is therefore the same as Carrrington's. Heliographic longitudes in the Carrington frame are then obtained by adding L0 to CMD. The final Carrington coordinates are stored in the columns Longitude and Latitude.

The Method field contains a character denoting the method by which the orientation of the solar disc was obtained. The most frequent value is '*C*' which stands for celestial system. The main horizontal line in the drawing was assumed to be parallel to the celestial equator. The orientation of the heliographic system is based on this assumption. The character '*Q*' stands for a rotational matching described in Section 2.2. The character '*H*' denotes observations without lines, for which we assumed that the orientation of the book is parallel to the horizon. If the observation was made at noon, this is equal to being parallel to the celestial equator. The apparent rotations of the disc drawn led us to the conclusion that discs at other times of the day are not oriented in a celestial, but rather in a horizontal system.

The Quality field gives a subjective quality of the positions on a scale from 1 to 3. All drawings with a pencil-drawn coordinate system obtained a Quality of 1. All drawings for which the Method is '*H*' obtained a Quality of 3. Drawings treated by rotational matching obtain a Quality of 1 for narrow probability distributions, a Quality of 2 for broad, but symmetric probability distributions, and a Quality of 3 for skewed probability distributions. Note that the quality flag only refers to the accuracy of the positions, not the spot sizes.

The values for Size are given from the original measurement. A conversion of these size classes into, e.g., microhemispheres is difficult and needs to be done at a later stage of comparing the Schwabe data with other sources. Spots were plotted as simple pencil dots of various size until 1831, while the first distinction between umbra and penumbra was made on 1831 January 06 and continued to be made throughout the rest of the observations.

Foreshortened spots near the solar limb were usually plotted as elliptical dots. In principle, our size estimates are projected areas; we tried to use a circular cursor shape which has an area equal to the elongated spot plotted. Given the difficulty in drawing arbitrarily thin lines with a pencil, however, we have to assume that these

Table 2. Data format of the data base of sunspot observations by Samuel Heinrich Schwabe for the period of 1825–1867. The fields are separated by one blank space each which is not included in the format declarations.

Field	Column	Format	Explanation
Year	1–4	I4	Year
Month	6–7	I2	Month
Day	9–10	I2	Day, referring to the German civil calendar running from midnight to midnight.
Hour	12–13	I2	Hour, times are mean local time.
Minute	15–16	I2	Minute, typically accurate to 15 min.
Timeflag	18	I1	Indicates how accurate the time is. Timeflag = 0 means the time has been inferred by the measurer (in most cases to be 12 h local time). Timeflag = 1 means the time is as given by the observer.
L0	20–24	F5.1	Heliographic longitude of apparent disc centre seen from Dessau.
B0	26–30	F5.1	Heliographic latitude of apparent disc centre seen from Dessau.
CMD	32–36	F5.1	Central meridian distance, difference in longitude from disc centre; contains -.- if line indicates spotless day; contains NaN if position of spot could not be measured.
Longitude	38–42	F5.1	Heliographic longitude in the Carrington rotation frame; contains -.- if line indicates spotless day; contains NaN if position of spot could not be measured.
Latitude	44–48	F5.1	Heliographic latitude, southern latitudes are negative; contains -.- if line indicates spotless day; contains NaN if position of spot could not be measured.
Method	50	C1	Method of determining the orientation. ‘*C*’: horizontal pencil line parallel to celestial equator; ‘*H*’: book aligned with azimuth elevation; ‘*Q*’: rotational matching with other drawings (spot used for the matching have ModelLong $\neq$ ‘−.−′, ModelLat $\neq$ ‘−.−′ and Sigma $\neq$ ‘−.−′).
Quality	52	I1	Subjective quality, all observations with coordinate system drawn by Schwabe get Quality = 1, also the ones with skewed systems that were rectified by the method described in Section 2.4. Positions derived from rotational matching may also obtain Quality = 2 or 3, if the probability distributions fixing the position angle of the drawing were not very sharp, or broad and asymmetric, respectively. Spotless days have Quality = 0; spots for which no position could be derived, but which have sizes, get Quality = 4.
Size	54–55	I2	Size estimate in 12 classes running from 1 to 12; a spotless day is indicated by 0.
SGroup	57–64	C8	Group designation taken from Schwabe.
Measurer	66–75	C10	Last name of person who obtained position.
ModelLong	77–81	F5.1	Model longitude from rotational matching (only spots used for the matching have this).
ModelLat	83–87	F5.1	Model latitude from rotational matching (only spots used for the matching have this).
Sigma	89–94	F6.3	Total residual of model positions compared with measurements of reference spots in rotational matching (only spots used for the matching have this). Holds for entire day.

projected areas are overestimated as compared to the spot sizes near the disc centre.

In case of days without sunspots, there is a single line in the data file with ‘-.-’ in the sunspot position, while we set Size = 0. Note that even then, we cannot provide a full record of Schwabe’s observations, since many of the 3699 verbal reports cannot be represented in this data format. The reports of spotless days are all incorporated in the data base with lines having Size = 0, while the remaining reports may be utilized in a future step of analysis of Schwabe’s observing records. There is usually only information on the appearance of new, or disappearance of existing groups, compared to the previous observation. Group sunspot numbers may easily be determined for these days, but only by assuming Schwabe’s definition of a group is correct (or compatible with our today’s understanding). It will also be possible to improve the group sunspot numbers by Hoyt & Schatten (1998) according to the verbal reports.

The column SGroup contains the group designation given by Schwabe. The Measurer column gives the last name of the person who obtained the spot position. The full names can be retrieved from the list of authors and the acknowledgements.

Additional information is given for the drawings that were analysed using the rotational matching. The spots used to fix the orientations of the drawings deliver posterior distributions for their positions as a side-product. We computed the averages of these posterior distributions and added the resulting positions to the corresponding lines in the data base as ModelLong and ModelLat. Since the model assumes stationary spots, the latitudes of the spots are constant for the drawings involved in this particular rotational matching. The longitudes are not exactly constant because they are Carrington longitudes, and the spots drift against the Carrington frame of reference according to the rotation profile (1) used. The Sigma column contains the standard deviation of the spots involved in the matching.

Occasionally, the model position does not refer to exactly the spot it is attached to. This results from spots that had split during the course of the period used for the rotational matching. The model position was then compared with the middle of the two new spots while the actual measurement afterwards, with the inferred orientation, generated two lines in the data base for the two spots.

4 SPOT DISTRIBUTION AND ACCURACY OF THE DRAWINGS

As already discussed in Section 2.2, the analyses of 75 drawings *without* reference lines delivered a plotting accuracy of 0.05 in units of the solar radius (2°.9 in the disc centre). We might consider this an upper limit, since the absence of reference lines makes accurate plotting more difficult, but see below.

The plotting accuracy certainly varied on a day-by-day scale, since poor weather may have allowed only little time for a drawing.

Table 3. Sample series for an estimate of the plotting accuracy.

Period	Days	σ
1832 Feb. 01–04	4	2°.13
1832 Feb. 10–15	4	3°.75
1832 Feb. 15–25	11	2°.56
1832 Feb. 24–Mar. 06	6	2°.54
1845 Jan. 12–24	7	4°.48
1845 Feb. 13–24	6	5°.06
1845 July 06–10	4	0°.91
1854 Apr. 06–14	7	2°.12
1855 Mar. 06–18	8	3°.08
1855 Oct. 23–Nov. 02	8	1°.46
1856 Apr. 10–20	9	1°.20
1864 Jan. 12–23	8	2°.07
1864 July 01–12	7	2°.95
1865 Feb. 05–14	5	3°.92
1865 July 09–19	8	1°.59
Simple average		2°.65

This is supported by occasional comments by Schwabe that the spot positions are only approximate because of clouds.

For the accuracy of the majority of drawings which do show a coordinate grid, we selected blindly a number of sequences of days during which simple spots of Waldmeier class H and I were crossing the solar disc. We determine the deviations of the measured latitudes from the average latitude of such a spot. To avoid problems with the differential rotation, we only looked at the scatter in the heliographic latitudes and assume that the true latitudes have not changed with time. The periods with the resulting standard deviations σ of the spots' latitudes are given in Table 3. The average σ (2°.65) needs to be converted into a total angular error, since we have only considered the latitudes here, whence approximately $\sigma_{\mathrm{tot}} \approx \sqrt{2}\sigma = 3^{\circ}.75$. Interestingly, this value is even a bit larger than the one obtained for the drawings without coordinate grids (2°.9). Note that proper motions in latitude are much smaller and extremely rarely exceed $0^{\circ}.1\,\mathrm{d}^{-1}$. They do not notably contribute to σ.

Fig. 4 shows the latitude–time distribution (butterfly diagram) of all sunspots measured in Schwabe's drawings. The patterns formed by the four cycles observed do not show any peculiarities at first glance. The separation of the two hemispheres is less distinct than in the butterfly diagram of the RGO/USAF data set. This is mostly due to the larger positional errors in the Schwabe data, and to a lesser extent due to the fact that the RGO/USAF data are average group positions while the Schwabe data contain individual spots which introduce an additional intrinsic scatter to the plot.

There were some periods in which the spot latitudes b were very high, $|b| > 50^{\circ}$. These were in 1836 August, when spot latitudes exceeded 60°, in 1839, in the middle of Cycle 8, when latitudes exceeded 50°, and in 1854 April when an individual spot was south of -50° at the end of Cycle 9. When inspecting the apparent motion of the spots across the disc, we noted that the coordinate system given in the drawings was not properly aligned. A total of 16 drawings have therefore been analysed using the rotational matching of Section 2.2. This method led to much lower latitudes for the first two periods mentioned. The matching of the last period in 1854 (a single spot over seven days) did not deliver sharp probability density distributions and was discarded. 1854 April 24 with the exceptional latitude was removed from the data base. Most of these problematic drawings were actually not made by Schwabe, but by other persons. The butterfly diagram also shows unusual latitudes in 1846 June. Inspection of the drawings shows, however, that the spot motion is consistent with the alignment of the drawings. We have not altered these measurements in the data base.

How likely are extreme latitudes? The RGO/USAF data contain minimum and maximum group latitudes of −59°.5 and 59°.7,

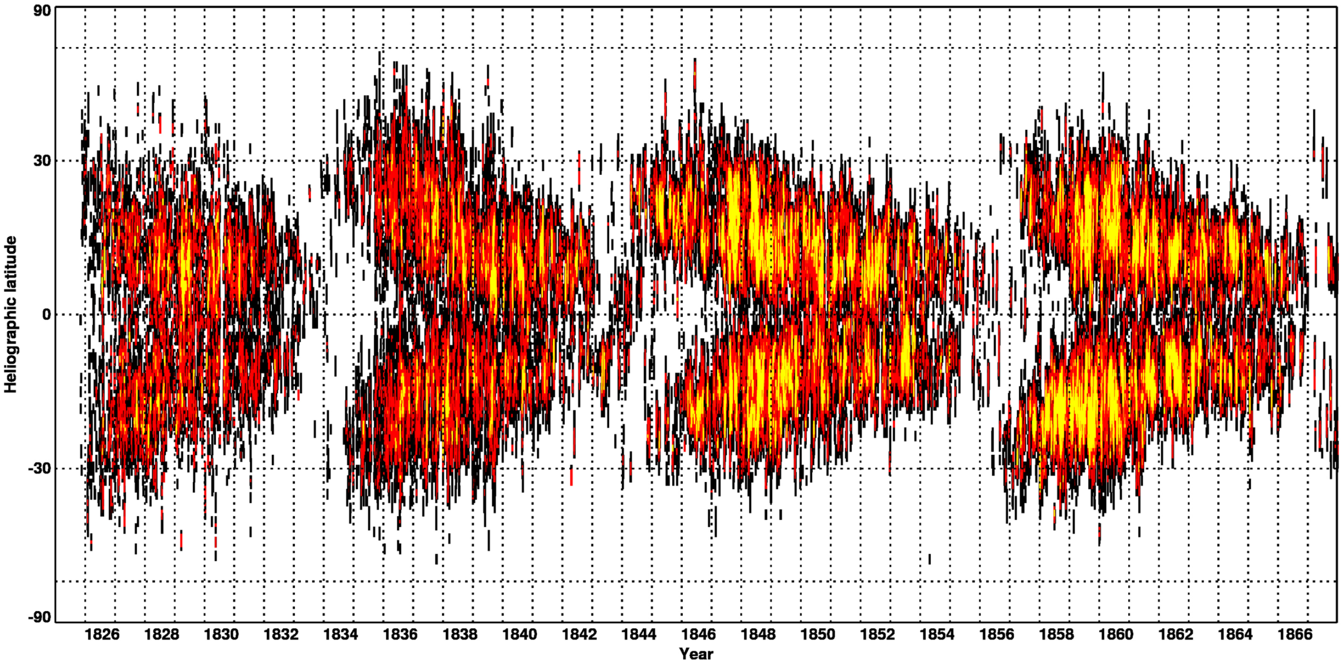

Figure 4. Butterfly diagram based on about 135 000 sunspot positions derived from Schwabe's observations of 1825–1867. A similar plotting style as used by Hathaway (http://solarscience.msfc.nasa.gov/SunspotCycle.shtml) is employed here.

respectively, according to the data base as of 2013 April 1.[3] Since these are average spot positions of a given group, the actual maximum and minimum latitudes of individual spots will be another few degrees towards the poles. A total of 14 sunspot groups have $|b| \geq 50^\circ$ in about 240 000 lines of data over almost 140 years in the RGO/USAF data base. The Schwabe measurements delivered 46 cases with $|b| \geq 50^\circ$ among about 135 000 lines of data, with extreme cases between $-52\overset{\circ}{.}8$ and $56\overset{\circ}{.}0$. There are relatively fewer high-latitude spots appearing in the RGO/USAF data than in Schwabe's data, but the extrema are comparable.

5 SUMMARY

We provide a set of about 135 000 sunspot positions and sizes measured on drawings by Samuel Heinrich Schwabe in the period of 1825 November 5 to 1867 December 29. The data base can be obtained from the website of the corresponding author.[4] The accuracy of the sunspot positions appears to be between 3° and 4° in the heliographic coordinate system near the disc centre. We also include all verbal reports on spotless days in the data base, so the file can also be used for studies of the activity. The data also contain an estimate of the individual spot sizes. They are given in 12 classes and should not be linearly scaled to physical areas.

The positions were obtained using (i) the coordinate system drawn by Schwabe, if available, (ii) a rotational matching with adjacent days if no coordinate system is given and (iii) an assumed alignment of the drawings with the horizontal system, if (i) and (ii) were not applicable, which was the case predominantly in the beginning of the observing period.

Note that we publish the first version of the data base here. The data file may be updated at some time in the future if errors emerge or the verbal information provides changes in the interpretation of the drawings (most likely concerning the clock times).

In the future, we intend to utilize also the information on spot evolution given in the verbal reports of Schwabe which are not accompanied by drawings. These improve the information on the lifetime of spots, since Schwabe carefully noted when spots disappeared and new spots appeared.

The potential of much less accurate drawings from the 18th century has been demonstrated by Arlt & Fröhlich (2012) who determined the differential rotation of the Sun based on the observations by Johann Staudacher. The more careful drawings by Schwabe will provide us with numerous quantitative results on four solar cycles in the 19th century.

ACKNOWLEDGEMENTS

RL thanks the Väisälä Foundation for financial support. NG thanks Deutsche Forschungsgemeinschaft for the support in grant Ar 355/7-1. We sincerely thank the Royal Astronomical Society for their permission to digitize and utilize the original manuscripts by Schwabe. We are very grateful to Anastasia Abdolvand, Luise Dathe, Jennifer Koch, Sophie Antonia Penger, Clara Ricken and Christianx Schmiel for their help with the measurements. We are highly indebted to Hans-Erich Fröhlich for providing his computer code for Bayesian parameter estimations.

REFERENCES

Arlt R., 2009, Sol. Phys., 255, 143
Arlt R., 2011, Astron. Nachr., 332, 805 (Paper I)
Arlt R., Fröhlich H.-E., 2012, A&A, 543, A7
Balthasar H., Vázquez M., Wöhl H., 1986, A&A, 155, 87
Cristo A., Vaquero J. M., Sánchez-Bajo F., 2011, J. Atmos. Sol.-Terr. Phys., 73, 187
Diercke A., Arlt R., Denker C., 2012, in Kosovichev A. G., de Gouveia Dal Pino E. M., Yan Y., eds, IAU Symp. 294, Cambridge Univ. Press, Cambridge, preprint (arXiv:1210.5856)
Hoyt D. V., Schatten K. H., 1998, Sol. Phys., 179, 189
Lefevre L., Clette F., 2012, Sol. Phys., doi:10.1007/s11207-012-0184-5
Lepshokov D. Kh., Tlatov A. G., Vasil'eva V. V., 2012, Geomagn. Aeron., 52, 843
Newcomb S., 1898, Tables of the Motion of the Earth on its Axis and Around the Sun. The Nautical Almanach Office, Washington

This paper has been typeset from a TEX/LATEX file prepared by the author.

---

[3] http://solarscience.msfc.nasa.gov/greenwch.shtml
[4] http://www.aip.de/Members/rarlt/sunspots